

Miti ali problemi dolgoročne digitalne hrambe?

UVOD

Hramba gradiva v digitalni obliki je tema, ki jo v praksi vedno pogosteje omenjamo. Največ govorimo o zajemu – digitalizaciji in hrambi tako zajetega gradiva. A veliki pok šele prihaja, saj vedno več sodobnega gradiva nastaja v elektronski obliki in bo prej ali slej začelo množično trkati na vrata naših arhivov. Organizacijske in tehnične priprave že tečejo, v prispevku pa pogledjmo nekaj vprašanj, ki se ob tem porajajo. Ali gre za resne probleme ali zgolj pretiran strah?

Tehnologija ni dorasla zahtevam e-hrambe

Sodobni elektronski sistemi, ki omogočajo nastajanje gradiva, so narejeni na osnovi najrazličnejših tehnologij, ki omogočajo izvajanje transakcij ter povezljivost sistemov. Ker bomo arhivirane podatke le še »brali«, za to ne bomo potrebovali vse zapletenosti tistih sistemov, v katerih so nastali. Velik poudarek bo namenjen skrbi za varnost in nespremenljivost podatkov. V velikih centrih lahko dosežemo potrebne ravni varnosti, da tudi dobro izurjen »zlikovec« težko naredi škodo. Za morebitno katastrofo pa v nasprotju s hrambo originalnega gradiva v papirni obliki s pridom izkoriščamo možnost za hkratno hrambo rezervnih kopij na rezervnih lokacijah. Z replikacijo tako strojne opreme kot podatkov znatno zmanjšamo verjetnost izgube gradiva (beri podatkov) in tako npr. ne okvara nosilca podatkov ne požar na kaki lokaciji s stališča hranjenih podatkov ne bosta povzročila škode. Še največjo nevarnost predstavlja človeški faktor: nevestno ali nekompetentno osebje, pomanjkanje pravil ali površno upoštevanje pravil. Zato potrebujemo stabilno in popolnoma obvladovano produkcijsko okolje. Vse poskusne dejavnosti ali celo improvizacije dovolimo le v testnih okoljih, ki so namenjena tudi za validacijo novih postopkov in šolanje osebja.

Podatek ali informacija: ali ne hranimo megle?

Informacijski sistemi hranijo in obdelujejo podatke. Razumevanje teh podatkov nas pripelje do informacij. Gole številke so podatki, a postanejo

informacije, če vemo na primer, da gre za meritve temperature vsak dan ob 12:00. Pojasnila o podatkih hranimo v metapodatkih.

Ker hranimo podatke, pričakujemo, da je shranjena informacija enaka originalni. Zato na primer ni potrebno hraniti originalnega CD-ja ali DVD-ja, saj tudi tak nosilec samo vsebuje vsebino, ki je bila nanj prepisana z drugega medija. Končni uporabnik mora imeti na voljo vpogled v vsebino podatkov (v informacije) v razumljivi obliki, enako kot bi dobil v roke popisan list papirja. Postopek hranjenja zato obsega tudi natančno opredelitev načina uporabe gradiva. Za ilustracijo naj bo najpreprostejši primer: v zvezi s številnimi datotekami v formatu PDF/A vemo, da bo njihova vsebina dostopna ob pomoči katerega od programov za prikaz vsebine PDF/A.

Kopije, kopije ... Katera je prava?

Na to vprašanje smo že odgovorili. Izkazalo se je, da so vse kopije v sistemu e-hrambe enakovredne, če le zagotovimo, da so enake. To pomeni, da je za vse potrebna enaka raven zaščite. Kljub temu si lahko pomagamo z »nadstandardno« zaščito za primer osrednjega trezorja s podatki, delovni podatki za pripravo oglednih kopij pa so v nekoliko dostopnejšem sistemu. Ob najmanjšem dvomu o avtentičnosti podatkov lahko podatke ponovno prepisemo iz trezorja.

Uporabnik moramo omogočiti »uporabo« hranjenega gradiva. Če mu ga posredujemo kot izpis na papirju ali datoteko, moramo poskrbeti za primerno opremljenost ali tehnično zaščito. Torej overitev ali elektronski podpis. Drugi dokumenti, ki krožijo po svetu, so nekontrolirane verzije in nikoli ne veš, ali ni že kaj spremenjeno.

Papir je lažje hraniti

Dolgoletna hramba na papirju je pripomogla k izkušnjam in dobremu poznavanju problematike. Rešiti pa moramo problem zaradi povečevanja količine sodobnega gradiva. Težko ocenimo, ali je v primerjavi z redkimi pisanji v preteklosti v vseh primerih takšna tudi njegova vrednost, a kar je namenjeno digitalno hrambi, bomo pač hranili. V veliki količini hranjenega gradiva se bo uporabnik naj-

bolje znašel s sodobnimi sredstvi, ki mu bodo omogočala iskanje ter obdelavo (da ne rečemo kar procesiranje) vsebin. Če papir ohranimo samo kot »rezervno kopijo za vsak primer« za morebitne katastrofe, se moramo vprašati, ali je prav papir najboljša oblika. Zanimiva alternativa je laserski zapis digitalne vsebine na mikrofilm, a v tem primeru potrebujemo računalnik, da pridemo do vsebine. Če bi se uresničil najbolj črn scenarij, lahko zanamcem pustimo samo mikrofilm in optične čitalce, ki ne potrebujejo elektrike. A ta tema (vključno s skrivnostnimi bunkerji v gorah) ni namen tega prispevka.

Prehitre odločitve danes se nam bodo maščevale

Z drugimi besedami, kar se bomo odločili zdaj, bomo zapečatili za naprej. In ker se vse spreminja, od verzij windows do pisarniških orodij, ta hip ni dobro hiteti ... A žal nas prav zato priganja čas. Kar se nam danes zdi domače, bo čez nekaj let eksotičen format. Sodobno hitenje k napredku, težnja po uporabi najnovejših tehnologij ali sodelovanje v »najnovejših« projektih zagotavljajo, da sprememb nikoli ne bo konec, pa naj pridejo po diktatu multinacionalk ali demokratičnih odprtokodnih združenj. To pa prinaša tveganje, da obstoječi formati v določenem trenutku ne bodo več ustrezni. Zato se že vnaprej pripravimo, da bodo spreminjanja formatov v nedoločenih časovnih intervalih potrebna vedno, ko bo tehnologija naredila preskok in bo lažje podpirala zase sodobne formate. Mogoče se motimo, a tu gre za preprost odziv na realno tveganje, kot ga zaznavamo zdaj.

Arhivisti smo zgodovinarji, kaj nam je tega treba?

Osnovna naloga arhivista se s sodobnimi oblikami gradiva ne spremeni. Delo še vedno zahteva tudi pomoč ustvarjalcem, odbiranje, popisovanje ipd. Res pa je, da »otip« (ogled) vsebine datoteke ne da enakega zadovoljstva kot pregledovanje starih arhivskih dokumentov. A s profesionalnim pristopom danes bomo omogočili, da bodo trenutno običajne vsebine ohranjene, da bodo lahko postale jutrišnja dragocenost. Potrebno je dodatno izobraževanje tako kot dandanes v vseh poklicih že zaradi interdisciplinarnega sodelovanja s strokovnjaki z drugih področij, ki nikakor niso le računalništvo in informatika. Večini od teh bodo prav arhivisti najboljše razložili, da zagotavljanje dolgoročne hrambe le ni nebodigatreba.

Dobro, a čemu to pretiravanje z notranjimi pravili?

Velika količina digitalnih podatkov predstavlja tveganje, da zaradi okvare ali pa kar z nepremišljenim preprostim ukazom računalniku nastane velikanska škoda. Ali pa, da drobna sprememba na enem od dokumentov nikoli ne bo razkrinkana. Tudi zaklenjen trezor ne pomaga, če iz njega štrli kup žic ... Kakorkoli že iščemo rešitve, bližnjic tu ni, razen kadar določeno gradivo razglasimo za malo pomembno ter zanemarimo verjetnost manipulacije. A to je daleč od zagotavljanja pravne varnosti. Sodobne tehnologije zaradi množične (robotizirane) izdelave omogočajo nekoč neslutene možnosti, vendar s tem nikakor ne nadomeščajo ustrezne organiziranosti poslovanja. Tega seveda ne moremo poenotiti, a prav v raznolikosti se skrivajo pasti in zato potrebujemo vsaj nekaj nadzora nad posameznimi rešitvami. Na prvo mesto lahko postavimo državno upravo – večina v njej pričakuje, »da bo že nekdo poskrbel za to«. A to smo mi.

Naš arhiv se bo spremenil v vzdrževalca starih aplikacij

Prevzem digitalnega gradiva od ustvarjalcev je neredko povezan z njihovo željo, da bi arhiv prevzel »vse«, torej tudi star program ter povrhu še obveznost, da bo zagotavljal določeno storitev ustvarjalca v identični obliki. Če bi šlo le za en program, bi se še pogovarjali. Natančen razmislek pokaže, da se v tem grmu skriva velik zajec – vsaka neuniverzalna rešitev (npr. prevzem v nestandardni obliki) zahteva veliko človeških virov in časa. Zato zagovarjamo špartanski pristop, pa čeprav bo dostop do podatkov nekoliko nerodnejši, kot je v originalni obliki danes. A vsaj bo.

PDF/A je rešitev vseh problemov

Format PDF/A omogoča hrambo vsebine dokumentov, pri tem pa je poudarek na prikazu vsebine. V tem se razlikuje od preostalih formatov, ki večinoma vzdržujejo notranji zapis vsebine, prikaz pa je rezultat obdelave modula za prikaz na ekranu ali modula za tiskanje. Pri tako hranjenem dokumentu bomo torej imeli zagotovljen enak prikaz v najrazličnejših okoljih. Ob množici najrazličnejših formatov in verzij je to dobra novica. A žal posamezni izdelki sodobnih orodij hranijo več informacij, kot je vidno na posameznem izpisu. Uporabnik v prihodnosti bi torej moral podatke prenesti v svoj sistem in jih tam obdelovati naprej.

Če pa to omogoča že sedanji format, bi lahko uporabniku prihranili nekaj dela. Zato pričakujemo, da se bodo počasi uveljavili tudi »pisarniški« formati, s tem pa bo omogočeno ohranjanje sodobnih vsebin, kadar bo to potrebno.

ZAKLJUČEK

V prispevku smo opisali nekaj žgočih dolgoročne hrambe v digitalni obliki. Ne gre za alternativo hrambi na papirju, saj govorimo le o gradivu, ki ga moramo hraniti v digitalni obliki. Re-

šitve so možne, pri tem pa se večina problemov ne skriva v tehnologiji, saj ta že dosega potrebno raven zrelosti in dostopnosti. Zaradi miselnega preskoka in odpiranja popolnoma novih vprašanj bo potrebnega kar nekaj časa ne le za iskanje vseh odgovorov, temveč tudi vseh vprašanj. Do boljšega razumevanja problematike nas bosta privedli stalno izobraževanje ter konkretno delo na tem področju, ki vključuje tudi spremljanje dogajanj v primerljivih institucijah v svetu in novih standardov.

Boris Domajnko



24. zborovanje Arhivskega društva Slovenije, Dolenjske Toplice, 14. do 16. oktober 2009